

UNCLASSIFIED

AD 432990

DEFENSE DOCUMENTATION CENTER

FOR

SCIENTIFIC AND TECHNICAL INFORMATION

CAMERON STATION, ALEXANDRIA, VIRGINIA



UNCLASSIFIED

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

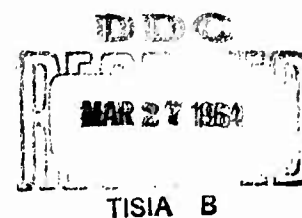
64-10

D1-82-0329

"Also available from the author"

BOEING SCIENTIFIC
RESEARCH
LABORATORIES

Maximum Likelihood Estimation for
Distributions with Monotone Failure Rate



Albert W. Marshall

Frank Proschan

Mathematics Research

February 1964

432990

EXTRACTED BY DDC
AS NO NO

432990

DL-82-0329

⑥ MAXIMUM LIKELIHOOD ESTIMATION FOR
DISTRIBUTIONS WITH MONOTONE FAILURE RATE

10 by
Albert W. Marshall
and
Frank Proschan

Mathematical Note No. 336
Mathematics Research Laboratory
BOEING SCIENTIFIC RESEARCH LABORATORIES
February 1964

Abstract

Using the idea of maximum likelihood, we derive an estimator for a distribution function possessing an increasing (decreasing) failure rate and also obtain corresponding estimators for the density and the failure rate. We show that these estimators are consistent.

Maximum Likelihood Estimation for Distributions
with Monotone Failure Rate

Albert W. Marshall and Frank Proschan
Boeing Scientific Research Laboratories

1. Introduction. Given a set of observations X_1, \dots, X_n from a common distribution function F , it is natural in the absence of additional information to estimate F by the usual empirical distribution function. However, one would not use this estimator if there were at hand sufficient a priori information about the distribution F , e.g., that F is a member of a given parametric class such as the normal. In this paper, we examine an intermediate case, the case that F is known to have increasing (decreasing) failure rate. Using the idea of maximum likelihood, we derive an estimator for F which itself has increasing (decreasing) failure rate, and also obtain estimators for the density and failure rate. These estimators are shown to be consistent.

2. Properties of IFR distributions, and formulation of the problem. The failure rate r of a distribution F having density f is defined by $r(x) = f(x)/[1 - F(x)]$, for $F(x) < 1$. It is easy to verify that if r is increasing, then $\int_0^x r(z)dz \equiv R(x) = -\log[1 - F(x)]$ is convex on the support of F , an interval. (Throughout this paper we write "increasing" for "nondecreasing" and "decreasing" for "nonincreasing.") Whether f exists or not, we say that F has increasing failure rate (IFR) if the

support of F is of the form $[\alpha, \beta]$ $-\infty \leq \alpha \leq \beta \leq \infty$, and if R is convex on $[\alpha, \beta]$. The importance of the IFR property and its applications to life testing and reliability are discussed in [3,4].

The continuous part of an IFR distribution F is absolutely continuous. To see this, choose $\epsilon > 0$, z such that $R(z) < \infty$, and points $\alpha_1 < \beta_1 < \alpha_2 < \beta_2 < \dots < \alpha_m < \beta_m \leq z$ satisfying $\sum_{i=1}^m (\beta_i - \alpha_i) < \epsilon/r^+(z)$, where $r^+(z) = \lim_{\delta \downarrow 0} [R(z+\delta) - R(z)]/\delta$ exists finitely since R is convex. Then

$$\sum_{i=1}^m |R(\beta_i) - R(\alpha_i)| = \sum_{i=1}^m \frac{R(\beta_i) - R(\alpha_i)}{\beta_i - \alpha_i} (\beta_i - \alpha_i) \leq r^+(z) \sum_{i=1}^m (\beta_i - \alpha_i) \leq \epsilon.$$

Thus R is absolutely continuous on $(-\infty, z)$, and the result follows.

For convenience, if F is IFR we define $r(x) = \infty$ for all x such that $F(x) = 1$. Note that for any distribution F and any x for which r is defined on $(-\infty, x)$, we have

$$(2.1) \quad 1 - F(x) = \exp\left[-\int_{-\infty}^x r(z) dz\right].$$

Further, properties of IFR distributions have been discussed in [2].

Let \mathcal{G} be the class of IFR distributions, and let $X_1 \leq X_2 \leq \dots \leq X_n$ be obtained by ordering a random sample from an unknown distribution F in \mathcal{G} . It is not possible to obtain a maximum likelihood estimator for $F \in \mathcal{G}$ directly by maximizing $\prod_{i=1}^n f(X_i)$, since for $F \in \mathcal{G}$, $f(X_n)$ can be arbitrarily large. Consequently, we first consider the subclass \mathcal{G}^M of distributions F in \mathcal{G} with corresponding failure rates bounded by M , obtaining

$\sup_{F \in \mathcal{G}^M} \prod_{i=1}^n f(X_i) \leq M^n$. We shall see that there is a unique distribution \hat{F}_n^M in \mathcal{G}^M

at which the supremum is attained. The conventional maximum likelihood estimators \hat{F}_n^M for \mathcal{G}^M converge in distribution as $M \rightarrow \infty$ (i.e., as $\mathcal{G}^M \rightarrow \mathcal{G}$) to an estimator $\hat{F}_n \in \mathcal{G}$ which we call maximum likelihood for \mathcal{G} . Furthermore, the density \hat{f}_n^M and failure rate \hat{r}_n^M of \hat{F}_n^M converge in a natural way to the density \hat{f}_n and failure rate \hat{r}_n (of the continuous part) of \hat{F}_n as is shown below in Section 3.

3. Derivation of the estimators. From (2.1) we obtain that the log likelihood $L = L(F)$ is given, for $F \in \mathcal{G}^M$, by

$$(3.1) \quad L = \sum_{i=1}^n \log r(X_i) - \sum_{i=1}^n \int_{-\infty}^{X_i} r(z) dz.$$

L is maximized over \mathcal{G}^M by a distribution with failure rate constant between observations, as may be seen as follows: Let $F \in \mathcal{G}^M$ have failure rate r and let F^* be the distribution with failure rate

$$(3.2) \quad r^*(x) = \begin{cases} 0, & x < X_1 \\ r(X_i), & X_i \leq x < X_{i+1}, i = 1, 2, \dots, n-1 \\ r(X_n), & x \geq X_n \end{cases}$$

Then $F^* \in \mathcal{G}^M$, and $r(x) \geq r^*(x)$ so that $-\int_{-\infty}^{X_i} r(z) dz \leq -\int_{-\infty}^{X_i} r^*(z) dz$ for all i ; we conclude that $L(F) \leq L(F^*)$. Thus, we may replace L by the function

$$(3.3) \quad \sum_{i=1}^n \log r(X_i) - \sum_{i=1}^{n-1} (n-i)(X_{i+1} - X_i)r(X_i).$$

The procedure for maximization of (3.3) subject to $r(X_1) \leq \dots \leq r(X_n) = M$ can be obtained as a direct application of [6, Corollary 2.1 and the discussion following] (see also [10, 11]). This procedure yields for r (corresponding

to $F \in \mathcal{G}^M$ the estimator

$$(3.4) \quad \hat{r}_n^M(X_i) = \min_{v \geq i+1} \max_{u \leq i} \left\{ \frac{1}{v-u} [r_u^{-1} + \dots + r_{v-1}^{-1}] \right\}^{-1}$$

where

$$(3.5) \quad r_j = [(n-j)(X_{j+1} - X_j)]^{-1}, \quad j = 1, 2, \dots, n-1, \quad r_n = M,$$

with the convention that $r_j = M$ when $[(n-j)(X_{j+1} - X_j)]^{-1} > M$.

The maximization procedure which yields (3.4) may be described as follows. First, find the maximum of (3.3) restricted only by $r(x) \leq M$, obtaining (3.5). If there is a reversal, say $r_i > r_{i+1}$, then set $r(X_i) = r(X_{i+1})$ in (3.3) and repeat the procedure. After at most n steps of this kind, a monotone estimator is obtained. The maximum derived with $r(X_i) = r(X_{i+1})$ can be directly obtained by replacing r_i and r_{i+1} by their harmonic mean, $(r_i^{-1} + r_{i+1}^{-1})^{-1}$. Succeeding steps amount to further such averaging which is extended just to the point necessary to eliminate all reversals. It can be seen that this is exactly what is called for in (3.4). (In this connection, see also [1,5].) The resulting estimator \hat{r}_n^M is of the form

$$\hat{r}_n^M(x) = \begin{cases} 0, & x < X_1 \\ r_{n_i+1, n_{i+1}}, & X_{n_i+1} \leq x < X_{n_{i+1}+1} \\ M, & x \geq X_n \end{cases}$$

where $r_{1, n_1} \leq r_{n_1+1, n_2} \leq \dots \leq r_{n_k+1, n-1}$, $0 = n_0 < n_1 < \dots < n_k < n-1$, and $r_{n_i+1, n_{i+1}}$ is the harmonic mean of $r_{n_i+1}, r_{n_i+2}, \dots, r_{n_{i+1}}$. Of course, the n_i are determined by the rule which determines the extent of

the averaging.

The estimator for r corresponding to $F \in \mathcal{G}$ is obtained by letting $M \rightarrow \infty$ in (3.4), and is given by

$$(3.6) \quad \hat{r}_n(X_i) = \min_{v \geq i+1} \max_{u \leq i} \left\{ \frac{1}{v-u} [(n-u)(X_{u+1} - X_u) + \dots + (n-v+1)(X_v - X_{v-1})] \right\}^{-1},$$

$i = 1, 2, \dots, n-1$ and $\hat{r}_n(X_n) = \infty$. For the remaining values of x , $\hat{r}_n(x)$ is determined by (3.2) with \hat{r}_n replacing r and r^* . The corresponding estimators \hat{F}_n and \hat{f}_n for F and f are obtained from \hat{r}_n using (2.1) and the relation $\hat{f}_n(x) = \hat{r}_n(x)[1 - \hat{F}_n(x)]$.

It is of interest to note that the estimator \hat{r}_n can also be written in the form

$$(3.7) \quad \hat{r}_n(x) = \inf_{v > x} \sup_{u < x} \left\{ \int_u^v [1 - F_n(y)] dy / [F_n(v) - F_n(u)] \right\}^{-1}$$

where F_n is the empirical distribution. Similarly, when $r(x)$ is increasing, it is given by (3.7) with F replacing F_n .

4. Consistency. In case we restrict ourselves to distributions with support contained in $[0, \infty)$, it is not difficult to verify that the regularity conditions used in [8] are satisfied by the family \mathcal{G}^M . Thus from the results of [8], it follows that $\hat{f}_n^M(t)$ is a consistent estimator of $f(t)$ for this restricted family. For fixed $t < \beta$, choose $M > r(t)$; then it follows that $\hat{f}_n^M(t) = \hat{f}_n(t)$. We conclude that $\hat{f}_n(t)$ is a consistent estimator of $f(t)$ for the family \mathcal{G} of IFR distributions F satisfying $F(0) = 0$. However, rather than verify the regularity conditions, we choose to give a direct proof of consistency. In so doing, we avoid the question of whether or not the regularity conditions are satisfied when $F(x) > 0$ for all x .

Theorem 4.1. If r is increasing, then for every t_0 ,

$$(4.1) \quad r(t_0-) \leq \liminf \hat{r}_n(t_0) \leq \limsup \hat{r}_n(t_0) \leq r(t_0+)$$

with probability one.

Proof. The right-hand inequality is trivial if $r(t_0+) = \infty$; otherwise, let $t_1 > t_0$ satisfy $r(t_1) < \infty$, and let $a_j(n) + 1$ be the index of the largest observation $\leq t_j$, $j = 0, 1$. Let $N_1(n)$ and $N_2(n)$ be defined by

$$\hat{r}_n(t_0) = \left[\frac{1}{N_2(n) - N_1(n)} \sum_{i=N_1(n)+1}^{N_2(n)} (n-i)(X_{i+1} - X_i) \right]^{-1}.$$

Let $Y = -[r(t_1)]^{-1} \log[1 - F(X)]$, so that $P\{Y > y\} = P\{1 - F(X) < e^{-r(t_1)y}\} = e^{-r(t_1)y}$, i.e., Y has an exponential distribution. Since the X_i are order statistics from the distribution F , $Y_i = -[r(t_1)]^{-1} \log[1 - F(X_i)]$ are order statistics from the exponential distribution, and $(n-i)(Y_{i+1} - Y_i)$ are independent, identically distributed exponential random variables, with mean $1/r(t_1)$. Finally,

$$(4.2) \quad Y_{i+1} - Y_i = [r(t_1)]^{-1} \left[\int_{-\infty}^{X_{i+1}} r(z) dz - \int_{-\infty}^{X_i} r(z) dz \right] = \int_{X_i}^{X_{i+1}} [r(z)/r(t_1)] dz \leq X_{i+1} - X_i, \quad i \leq a_1(n).$$

From (3.6) and (4.2), it follows that

$$\begin{aligned} \hat{r}_n(t_0) &\leq \left[\frac{1}{a_1(n) - N_1(n)} \sum_{i=N_1(n)+1}^{a_1(n)} (n-i)(X_{i+1} - X_i) \right]^{-1} \\ &\leq \left[\frac{1}{a_1(n) - N_1(n)} \sum_{i=N_1(n)+1}^{a_1(n)} (n-i)(Y_{i+1} - Y_i) \right]^{-1}. \end{aligned}$$

But

$$\lim_{n \rightarrow \infty} \left[\frac{1}{a_1(n) - N_1(n)} \sum_{i=N_1(n)+1}^{a_1(n)} (n-i)(Y_{i+1} - Y_i) \right]^{-1} = r(t_1)$$

with probability one, by the strong law of large numbers

($\lim_n [a_1(n) - N_1(n)] \geq \lim_n [a_1(n) - a_0(n)] = \infty$ with probability one).

We conclude that $\limsup_n \hat{r}_n(t_0) \leq r(t_1)$ with probability one, and the right-hand inequality of (4.1) follows. A similar proof yields the left-hand inequality. ||

Corollary 4.2. If r is increasing, then for all t , $\lim_{n \rightarrow \infty} \hat{F}_n(t) = F(t)$ with probability one.

Proof. It is sufficient to prove the theorem for t satisfying $F(t) < 1$, in which case $\hat{F}_n(t) < 1$ for sufficiently large n . By Theorem 4.1, $\lim_{n \rightarrow \infty} \hat{r}_n(z) = r(z)$ except possibly for z in a set of Lebesgue measure zero. For $z \in [x, t]$, $x > -\infty$, $\hat{r}_n(z) < \infty$, and by the Lebesgue dominated convergence theorem, $\lim_{n \rightarrow \infty} \int_x^t \hat{r}_n(z) dz = \int_x^t r(z) dz$ with probability one. Then, by (2.1),

$$(4.3) \quad \lim_{n \rightarrow \infty} \frac{1 - \hat{F}_n(t)}{1 - \hat{F}_n(x)} = \frac{1 - F(t)}{1 - F(x)} \quad \text{with probability one.}$$

If we knew that $F(x) = 0$ for some $x > -\infty$, this would complete the proof.

In order to obtain an upper bound for $\int_{-\infty}^x \hat{r}_n(z) dz$, we first note that

$$\begin{aligned} r_{n_{i+1}, n_{i+1}} &= \left\{ \frac{1}{(n_{i+1}+1) - (n_i+1)} [(n - (n_i+1))(X_{n_{i+2}}) + \dots + (n - n_{i+1})(X_{n_{i+1}+1} - X_{n_{i+1}})] \right\}^{-1} \\ &\leq \left\{ \frac{1}{(n_{i+1}+1) - (n_i+1)} [(n - n_{i+1})(X_{n_{i+2}} - X_{n_{i+1}}) + \dots + (n - n_{i+1})(X_{n_{i+1}+1} - X_{n_{i+1}})] \right\}^{-1} \\ &= \frac{n_{i+1} - n_i}{(n - n_{i+1})(X_{n_{i+1}} - X_{n_{i+1}+1})} \end{aligned}$$

Let $k \equiv k(n)$ be the index of the largest observation not greater than x ;
if X_k is in $[X_{n_j+1}, X_{n_{j+1}+1})$, we obtain by (3.6),

$$\begin{aligned}\hat{r}_{n_j+1, n_{j+1}} &\leq \left\{ \frac{1}{(k+1)-(n_j+1)} [(n-(n_j+1))(X_{n_j+2} - X_{n_j+1}) + \dots + (n-k)(X_{k+1} - X_k)] \right\}^{-1} \\ &\leq \left\{ \frac{1}{k-n_j} [(n-k)(X_{k+1} - X_{n_j+1})] \right\}^{-1}.\end{aligned}$$

From these estimates, it follows that

$$\int_{-\infty}^x \hat{r}_n(z) dz \leq \sum_{i=0}^{j-1} \hat{r}_{n_i+1, n_{i+1}} (X_{n_{i+1}+1} - X_{n_i+1}) + \hat{r}_{n_j+1, n_{j+1}} (X_{k+1} - X_{n_j+1}) \leq \frac{k}{n-k}.$$

If $0 < \epsilon < \frac{1}{2}$ and x satisfies $F(x) \leq \epsilon$, then $\lim \frac{k}{n-k} = \frac{F(x)}{1-F(x)} < \frac{\epsilon}{1-\epsilon} < 2\epsilon$ with probability one, so that $\limsup \int_{-\infty}^x \hat{r}_n(z) dz < 2\epsilon$, and by (2.1), $\liminf [1 - \hat{F}_n(x)] \geq e^{-2\epsilon} \geq 1 - 2\epsilon$. This together with (4.3) completes the proof. ||

Corollary 4.3. If r is increasing and continuous on $[a, b]$, then

- (i) $\lim_{n \rightarrow \infty} \sup_{t \in [a, b]} |\hat{r}_n(t) - r(t)| = 0$
- (ii) $\lim_{n \rightarrow \infty} \sup_{-\infty < t < \infty} |\hat{F}_n(t) - F(t)| = 0,$
- (iii) $\lim_{n \rightarrow \infty} \sup_{t \in [a, b]} |\hat{f}_n(t) - f(t)| = 0,$

each with probability one.

Proof. (i) and (ii) follow from the same methods as in the usual proof of the Glivenko-Cantelli theorem. (iii) follows from (i), (ii), and the fact that $f(t) = r(t)[1 - F(t)]$. ||

5. Comparison between $r_n(t)$ and $\hat{r}_n(t)$. We shall show that with respect to a certain metric $\hat{r}_n(t)$ is closer to $r(t)$ than is $r_n(t)$,

where

$$(5.1) \quad r_n(t) = \begin{cases} 0 & \text{for } 0 \leq t < X_1 \\ \{(n-j)(X_{j+1}-X_j)\}^{-1} & \text{for } X_j \leq t < X_{j+1}, j = 1, 2, \dots, n-1 \\ \infty & \text{for } X_n \leq t < \infty. \end{cases}$$

Note that $r_n(t)$ represents the "unaveraged" estimate of the failure rate, i.e., the estimate that does not take into account the requirement that $r(t)$ be increasing. The result is similar to an inequality of [1, page 644] and is really a special case of the results of [7]. We give a simple proof for convenience and completeness.

We need the general result:

Theorem 5.1. Let h be nondecreasing, g be integrable with respect to the measure μ , the discontinuities of h distinct from the points at which μ places positive mass, and

$$\tilde{g}(x) = \sup_{s \leq x} \inf_{t \geq x} \frac{\int_s^t g(\theta) d\mu(\theta)}{\mu(t) - \mu(s)}.$$

Then

$$(5.2) \quad \int (g-h)^2 d\mu \geq \int (\tilde{g}-h)^2 d\mu + \int (g-\tilde{g})^2 d\mu.$$

Proof. It suffices to show $\int (\tilde{g}-h)(g-\tilde{g}) d\mu \geq 0$. The x -axis can be broken up into single points and maximal intervals on each of which $\tilde{g}(x)$ is constant. At a single point x , $\tilde{g}(x) = g(x)$. Let $[a, b]$ be an interval, with $\tilde{g}(x) = \tilde{g}$ on $[a, b]$. Define $G(x) = \int_{-\infty}^x g(\theta) d\mu(\theta)$. Then

$$\int_a^b (\tilde{g} - h(x))(g(x) - \tilde{g})d\mu(x) = \int_a^b [G(x) - G(a) - \tilde{g}[\mu(x) - \mu(a)]]d[h(x) - \tilde{g}] \geq 0,$$

$$\text{since } \frac{G(x) - G(a)}{\mu(x) - \mu(a)} \geq \inf_{t \geq x} \frac{G(t) - G(a)}{\mu(t) - \mu(a)} = \sup_{s \leq x} \inf_{t \geq x} \frac{G(t) - G(s)}{\mu(t) - \mu(s)} = \tilde{g}. \parallel$$

Identifying $h(t)$ as $r(t)$, $g(t)$ as $r_n(t)$, $\tilde{g}(t)$ as $\hat{r}_n(t)$, and $\mu(-\infty, t]$ as $F_n(t)$, the usual empirical distribution, we obtain from Theorem 5.1:

Theorem 5.2. With probability one,

$$(5.3) \quad \int_{-\infty}^{\beta^-} \{r_n(t) - r(t)\}^2 dF_n(t) \geq \int_{-\infty}^{\beta^-} \{\hat{r}_n(t) - r(t)\}^2 dF_n(t) + \int_{-\infty}^{\beta^-} \{r_n(t) - \hat{r}_n(t)\}^2 dF_n(t).$$

Thus, in the sense made precise by (5.3), $\hat{r}_n(t)$ is closer to $r(t)$ than is $r_n(t)$.

6. Decreasing failure rate. A distribution F is said to have decreasing failure rate (DFR) if the support of F is of the form $[\alpha, \infty)$, $\alpha > -\infty$, and if $\log[1 - F(x)]$ is convex on $[\alpha, \infty)$. Such distributions arise, e.g., as mixtures of exponentials (see [9]).

If F is DFR then by an argument similar to that used in the IFR case, it is absolutely continuous except possibly for a discontinuity at the point α . Thus, the measure determined by F is absolutely continuous with respect to $\mu_\alpha = \delta_\alpha + \lambda$ where δ_α places unit mass on $\{\alpha\}$ and λ is Lebesgue measure; we denote the density of F with respect to μ_α by f , and again define the failure rate of F by $r(x) = f(x)/[1 - F(x-)]$. If F is DFR, we always take a version of f for which r is decreasing in (α, ∞) .

Allowing for the fact that f is a density with respect to μ_α , we see that (2.1) is replaced by

$$(6.1) \quad 1 - F(x) = [1 - r(\alpha)] \exp\left[- \int_{(\alpha, x]} r(z) dz\right].$$

Estimation in the DFR case parallel that in the IFR case, but with some interesting differences. The first of these is that there are really two problems in the DFR case, depending on whether or not the point α is known.

First consider the case that α is known and suppose $\alpha = X_1 = \dots = X_k < X_{k+1} < \dots < X_n$ (in case $k = 0$, we define $X_0 = \alpha$). Using (6.1), $f(x) = r(x)[1 - F(x-)]$ and the relations $r(\alpha) = f(\alpha) = F(\alpha+)$, we write the log likelihood in the form

$$k \log r(\alpha) + (n-k) \log (1 - r(\alpha)) + \sum_{i=k+1}^n \log r(X_i) - \sum_{i=k+1}^n \int_{\alpha}^{X_i} r(z) dz.$$

Maximization of the first two terms yields $\hat{r}_n(\alpha) = k/n = \hat{F}(\alpha+)$. Maximization of the last two terms is quite analogous to that in the IFR case, and yields for r the estimator

$$\hat{r}_n(x) = \hat{r}_n(X_i), \quad X_{i-1} < x \leq X_i, \quad i = k+1, \dots, n$$

where

$$\hat{r}_n(X_i) = \max_{v \geq i} \min_{u \leq i-1} \left\{ \frac{1}{v-u} [(n-u)(X_{u+1} - X_u) + \dots + (n-v+1)(X_v - X_{v-1})] \right\}^{-1},$$

and $X_0 = \alpha$ in case $k = 0$.

Contrary to the IFR case, this DFR estimator is not unique; it is determined by the likelihood equation only for $x \leq X_n$, and may be extended beyond X_n in any manner that preserves the DFR property.

Consider now the case that α is unknown, and assume for the moment that F is absolutely continuous with respect to Lebesgue measure. If F is DFR on $[\alpha, \infty)$ for $\alpha > X_1$, then the likelihood $\Lambda(F) = \prod f(X_i) = 0$. If F is DFR on $[\alpha, \infty)$ for $\alpha < X_1$, then $\Lambda(F) < \Lambda(\tilde{F})$ where \tilde{F} is defined by

$$\tilde{F}(x) = \begin{cases} \frac{F(x) - F(X_1)}{1 - F(X_1)}, & x \geq X_1 \\ 0 & x < X_1. \end{cases}$$

Thus the maximum likelihood estimator for α unknown is found among those DFR distributions with support $[X_1, \infty)$, and the problem reduces to the case of known α . Note that the estimator \hat{F}_n has a jump of at least $1/n$ at X_1 .

The proof of consistency in the DFR case is similar to the proof in the IFR case.

7. The discrete case. A related problem of interest occurs in the case that F is discrete IFR. If F is a discrete distribution with mass p_i at x_i , $i = \dots, -1, 0, 1, 2, \dots$ and the x_i are ordered increasingly, the ratio

$$\rho_i = p_i / \sum_{j=i}^{\infty} p_j, \quad i = \dots, -1, 0, 1, 2, \dots,$$

is called the (discrete) failure rate of F . If r_i is increasing, then F is said to be discrete IFR. It is easily verified that

$$p_i = \rho_i \prod_{j=-\infty}^{i-1} (1 - \rho_j), \quad i = \dots, -1, 0, 1, \dots$$

If a sample of n independent observations from F consists of m_i occurrences at x_i , where for notational convenience, $i = 1, 2, \dots, k$, then the log likelihood function is

$$L = \sum_{i=1}^k m_i \log p_i = \sum_{i=1}^k \{m_i \log p_i + (m_{i+1} + \dots + m_k) \log(1 - p_i)\}.$$

We wish to maximize L subject to $p_1 \leq p_2 \leq \dots \leq p_k$.

With proper identification, this problem is exactly the one solved in [1]. The solution is obtained by averaging (through adding numerators and denominators) the quantities

$$(7.1) \quad \rho_i^* = \begin{cases} 0, & i < 1 \\ m_i / (m_i + \dots + m_k), & i = 1, 2, \dots, k \\ 1, & i > k, \end{cases}$$

to eliminate any reversals $\rho_j^* > \rho_{j+1}^*$. After sufficient averaging a set of increasing estimates $\hat{\rho}_1, \dots, \hat{\rho}_k$ are obtained which may be written as

$$(7.2) \quad \hat{\rho}_i = \min_{k \geq r \geq i} \max_{s \leq i} \frac{m_s + m_{s+1} + \dots + m_r}{\sum_{j=r}^s (m_j + \dots + m_k)}.$$

The estimator given in Section 3 for the continuous case may be derived from this as a limiting case. Consistency of the estimator (7.2) follows as in [1].

If ρ_i is decreasing, then F is said to be discrete DFR. In this case, maximum likelihood estimators may be obtained and consistency proved using the same method as in the discrete IFR case with obvious modifications.

Acknowledgment. We would like to thank Professor Ronald Pyke for his helpful advice and suggestions.

References

- [1] Ayer, Miriam, Brunk, H. D., Ewing, G. M., Reid, W. T., and Silverman, Edward (1955). An empirical distribution function for sampling with incomplete information. Ann. Math. Statist. 26, 641-647.
- [2] Barlow, R. E., Marshall, A. W., and Proschan, F. (1963). Properties of probability distributions with monotone hazard rate. Ann. Math. Statist. 34, 375-389.
- [3] Barlow, R. E. and Proschan, F. (1962). Planned replacement, appearing in Studies in Applied Probability and Management Science, edited by Arrow, Karlin, and Scarf, Stanford University Press, Stanford, California.
- [4] Barlow, R. E. and Proschan, F. (1963). Comparison of replacement policies, and renewal theory implications. Accepted for publication in Ann. Math. Statist., Boeing document DL-82-0237, February 1963.
- [5] Brunk, H. D. (1955). Maximum likelihood estimates of monotone parameters. Ann. Math. Statist. 26, 607-616.
- [6] Brunk, H. D. (1958). On the estimation of parameters restricted by inequalities. Ann. Math. Statist. 29, 437-454.
- [7] Brunk, H. D. (1961). Best fit to a random variable by a random variable measurable with respect to a σ -lattice. Pacific J. Math. 11, 785-802.
- [8] Crawford, Gordon and Saunders, S. C. (1964). Nonparametric maximum likelihood estimation. Ann. Math. Statist., accepted for publication, Boeing document DL-82-0308, October 1963.
- [9] Proschan, F. (1963). Theoretical explanation of observed decreasing failure rate. Technometrics 5, 375-383.
- [10] Van Eeden, C. (1956). Maximum likelihood estimation of ordered probabilities, Indag. Math. 18, 444-455.

- [11] Van Eeden, C. (1957). Maximum likelihood estimation of partially or completely ordered parameters. Indag. Math. 19, 128-136 and 201-211.